Practical methods to pool multi-study joint longitudinal and time to event data

Maria Sudell, Dr Ruwanthi Kolamunnage-Dona, Dr Catrin Tudur Smith

University of Liverpool

mesudell@liverpool.ac.uk

Aims of Talk

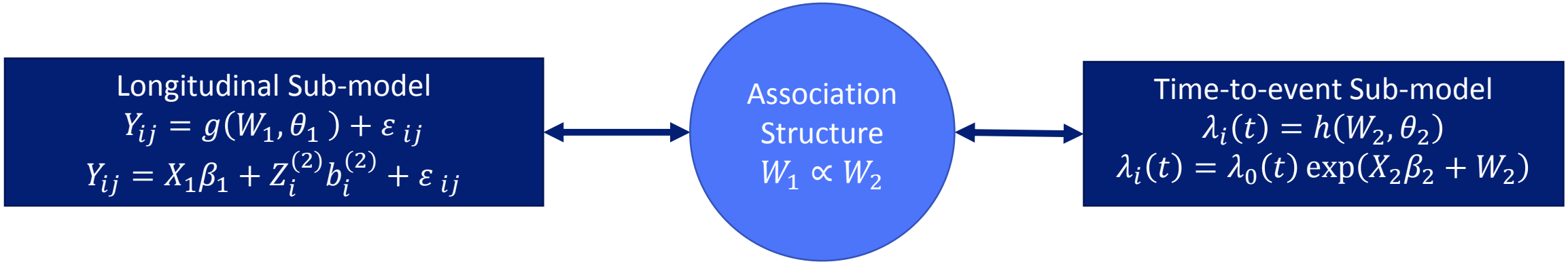
- Overview of methodology for
  - Two stage meta-analysis of multi-study joint longitudinal and time-to-event data
  - One stage meta-analysis of multi-study joint longitudinal and time-to-event data

- Review of current reporting of single study joint models applied to medical datasets

- Introduction of software package in R to implement methods

Joint longitudinal and time-to-event data (single study)

- Methods to simultaneously model potentially related longitudinal and time-to-event data
- Can produce less biased more efficient results than standalone cases where linked longitudinal and time-to-event data exists

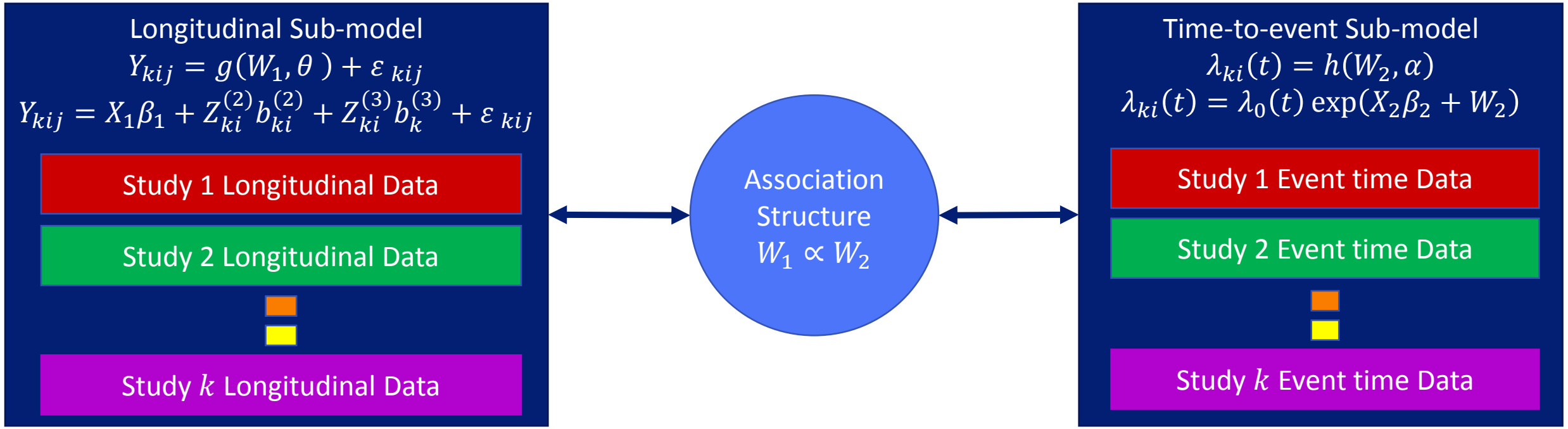


Meta-Analysis (MA)

- Systematic pooling of results from multiple studies
- Allows increased precision, identification of effect sizes too small to be identified in single studies, and allows questions additional to those originally posed in the data to be answered
- Gold standard – Individual Participant/Patient Data (IPD) meta-analyses, where data for each individual recorded in studies identified in the meta-analysis is available.

Joint longitudinal and time-to-event data (multi-study)

- Data available from multiple studies
- Clustering of data within studies must be accounted for (e.g. through random effects, interaction terms, stratified baseline hazard)



Approaches to modelling multi-study IPD joint data

- Two main approaches – one stage or two stage

- Two stage approaches
  - Separate joint models fitted to data from each study
  - Results from each study pooled using standard meta-analytic techniques

- One stage approaches
  - Joint model fitted to meta-dataset (containing data from all studies)
  - Clustering of data must be accounted for

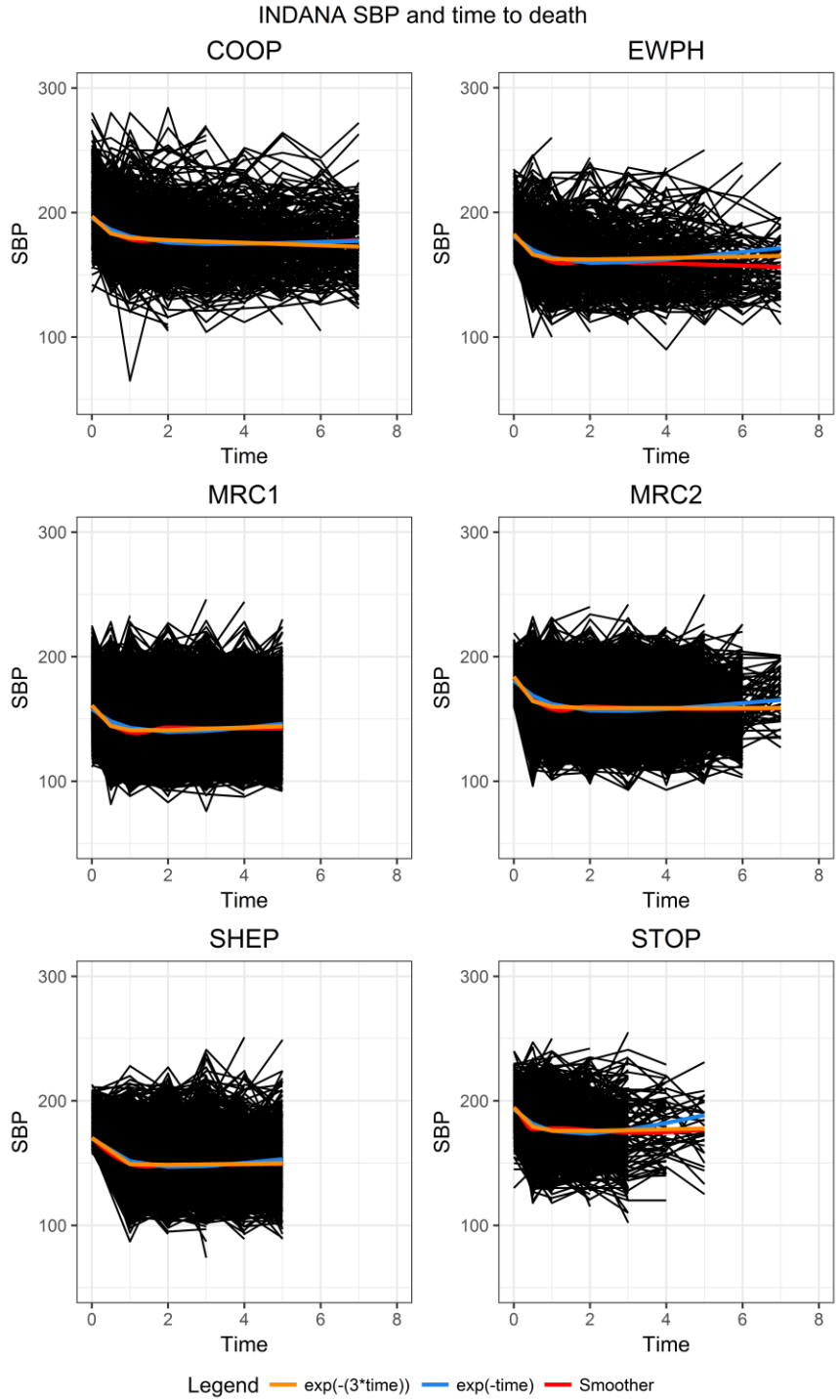
Real Data – subset of the INDANA dataset

- IPD from multiple studies investigating the effect of no treatment versus any treatment for hypertensive patients

- Longitudinal data measured at baseline, 6 months, then annually thereafter to maximum of 7 years. Measurement patterns varied between studies

- Using subset with data for longitudinal outcome systolic blood pressure (SBP) and time to death, data available from 6 studies. Proportions of individuals from each study, and proportions events/censored within each study kept same as full dataset. Full analysis using entire dataset currently running.

- Evidence of a changepoint in the data at 6 month, so $\exp(-3 * time)$ term included in the model


INDANA SBP and time to death
COOP
EWPH
MRC1
MRC2
SHEP
STOP
SBP
Time
300
200
100
0
2
4
6
8
Legend
exp(-(3*time))
exp(-time)
Smoother

Two stage methods

Two stage methods - overview

Stage 1: Joint model fitted to data from each study

Longitudinal Sub-model

$$Y_{kij} = \beta_{10} + \beta_{11} time_{kij} + \beta_{12} \exp(-3 * time_{kij}) + \beta_{13} treat_{ki} + b_{0ki}^{(2)} + b_{1ki}^{(2)} time_{kij} + \varepsilon_{kij}$$

Association Structure

$$W_{ki}^{(2)} = \alpha^{(2)}(b_{0ki}^{(2)} + b_{1ki}^{(2)} time_{kij})$$

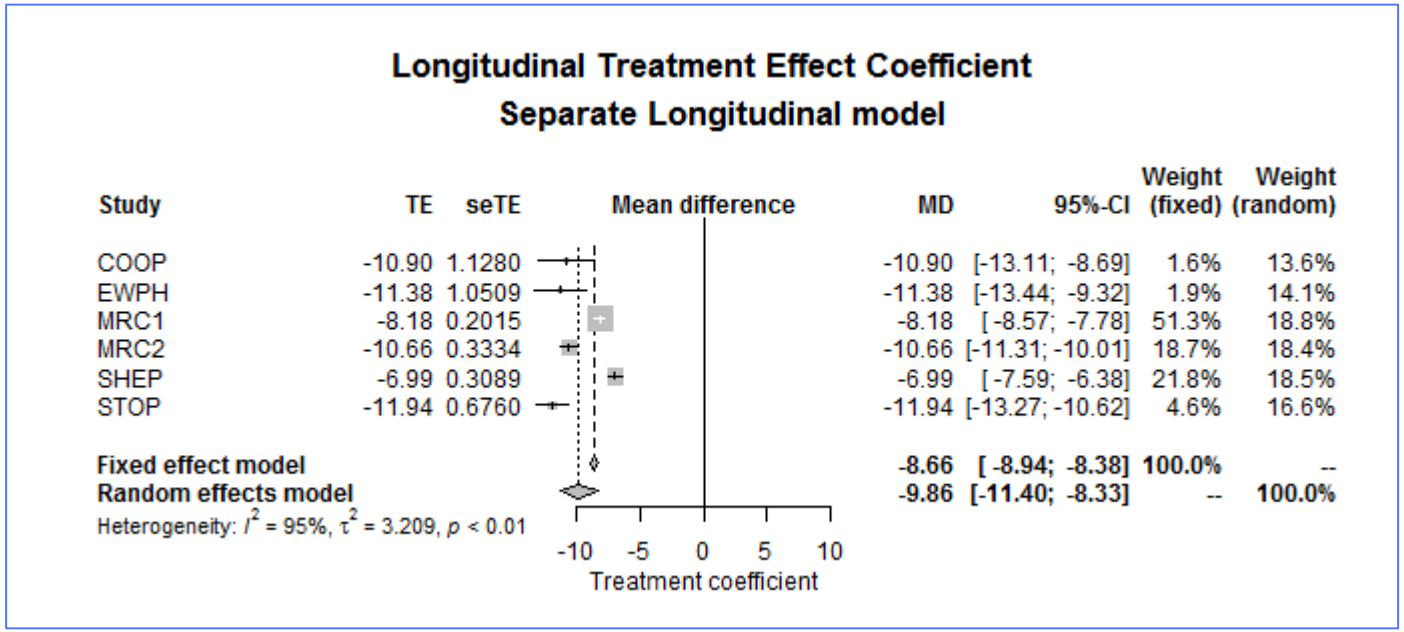
Time-to-event Sub-model

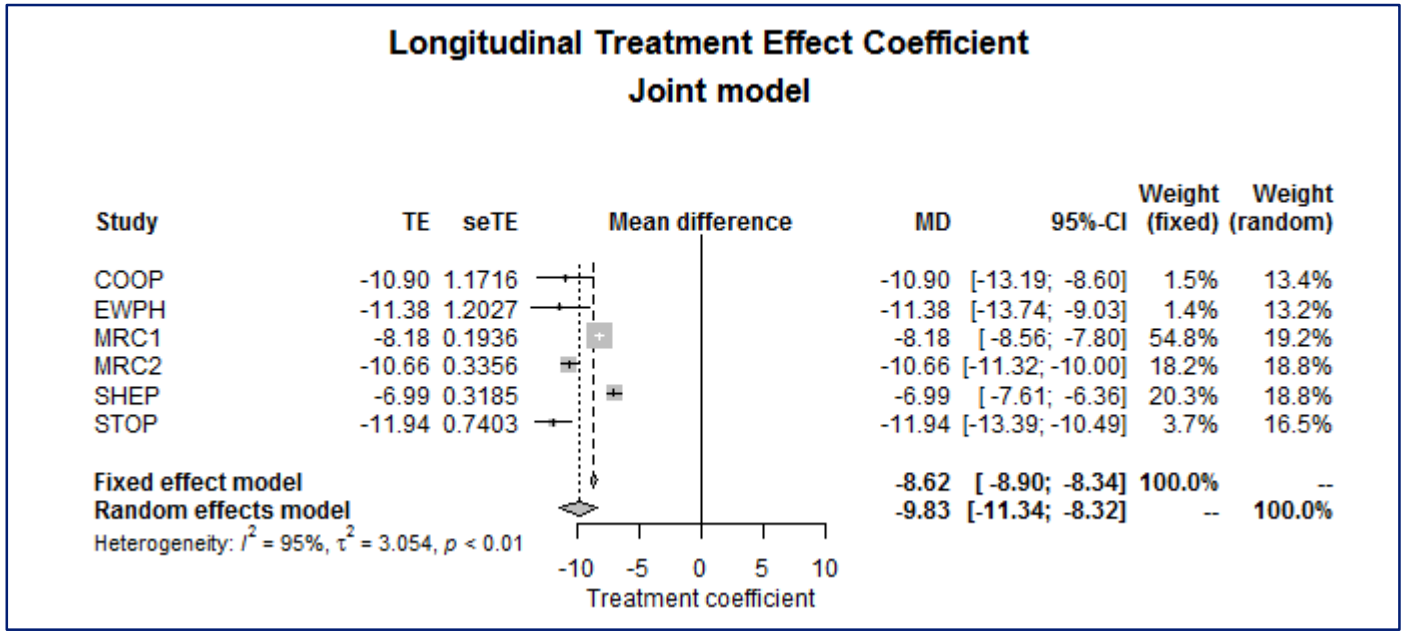
$$\lambda_{ki}(t) = \lambda_0(t) \exp(\beta_{21} treat_{ki} + W_{ki}^{(2)})$$

Stage 2: Study specific parameters pooled using standard meta-analytic techniques

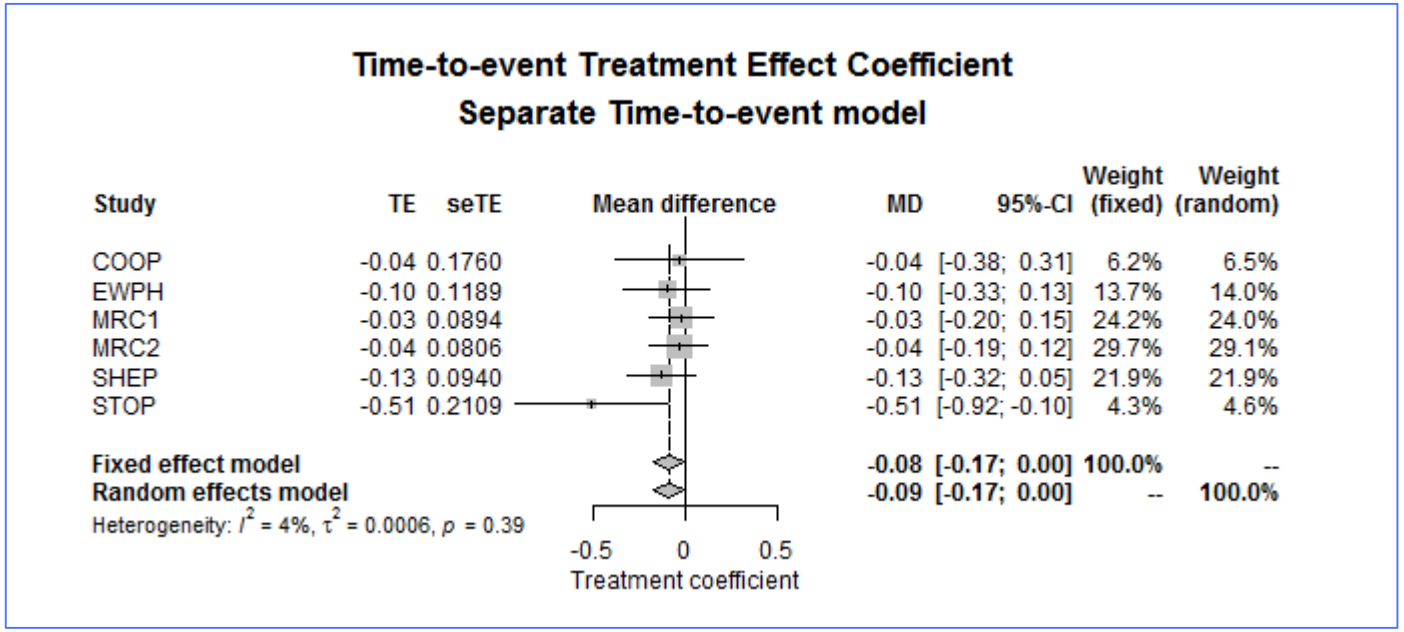
- Inverse variance method used (DerSimonian method used for random meta-analyses)
- Both fixed and random effects meta-analyses fitted and compared
- Separate meta-analyses for each parameter of interest

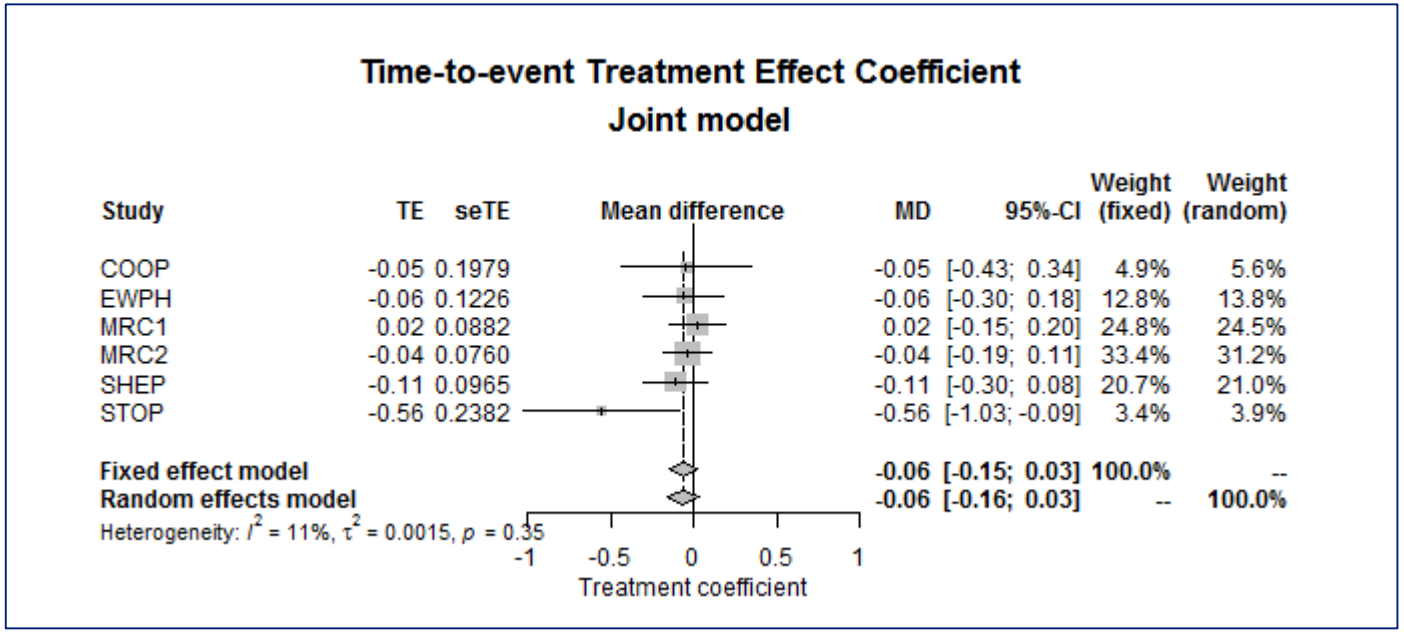
Two stage methods – real data



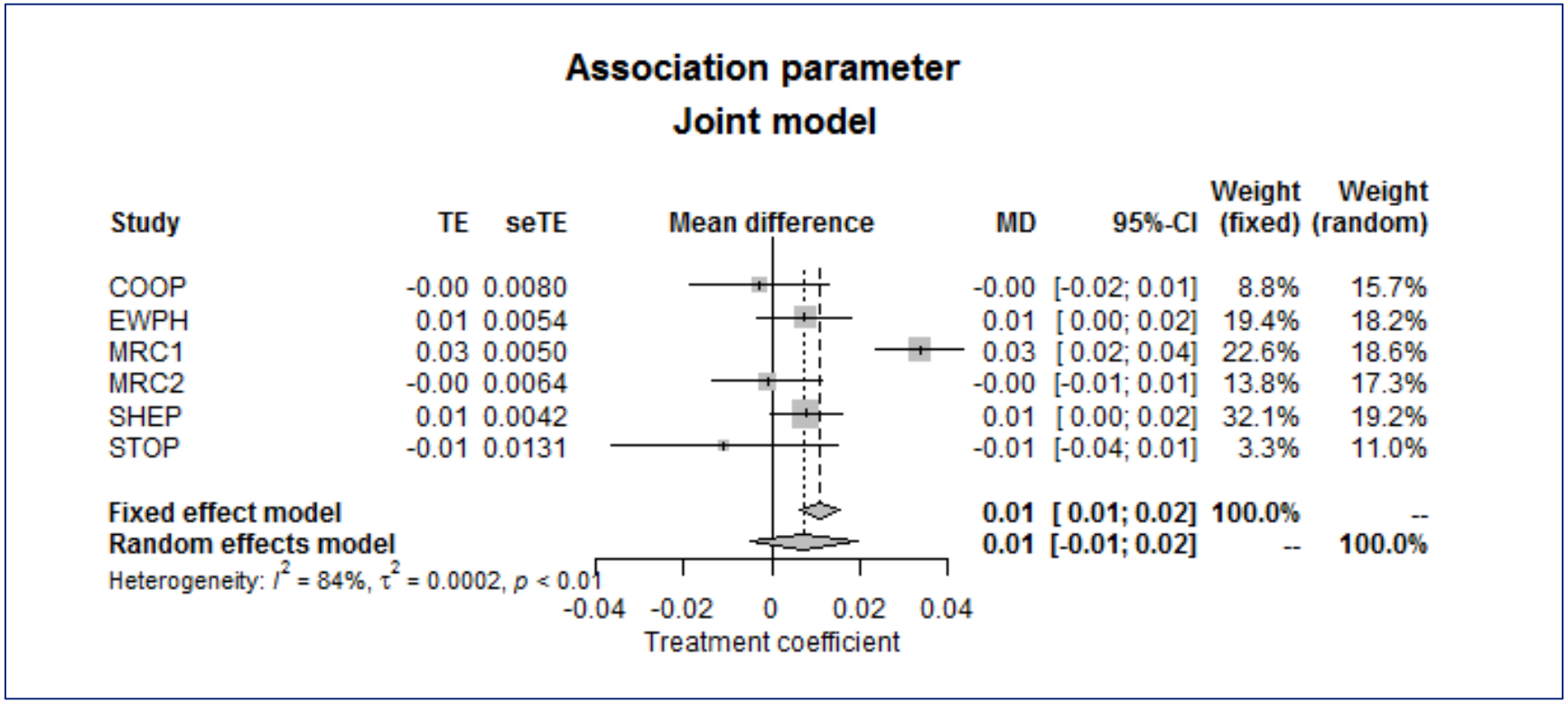


Two stage methods – real data

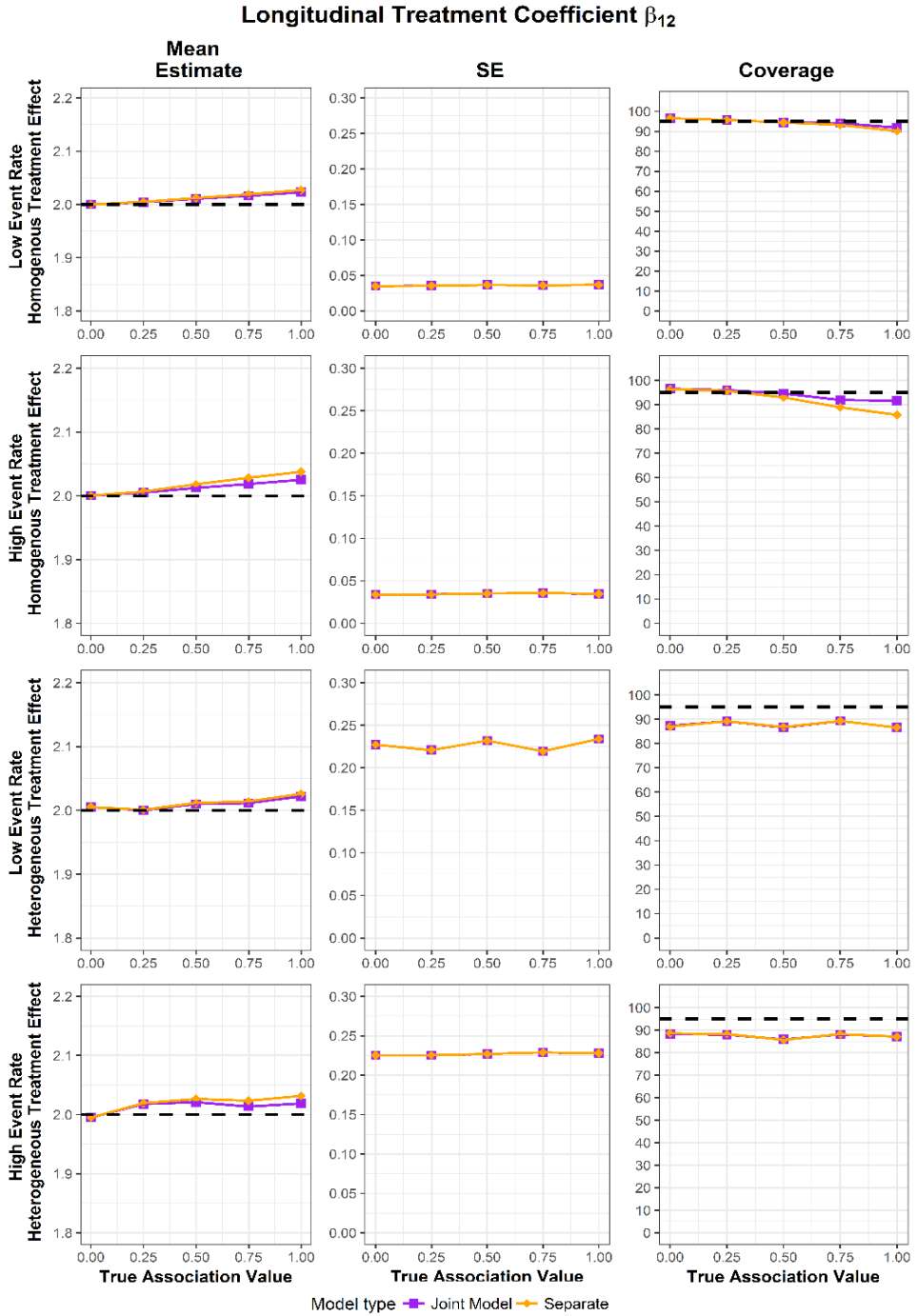




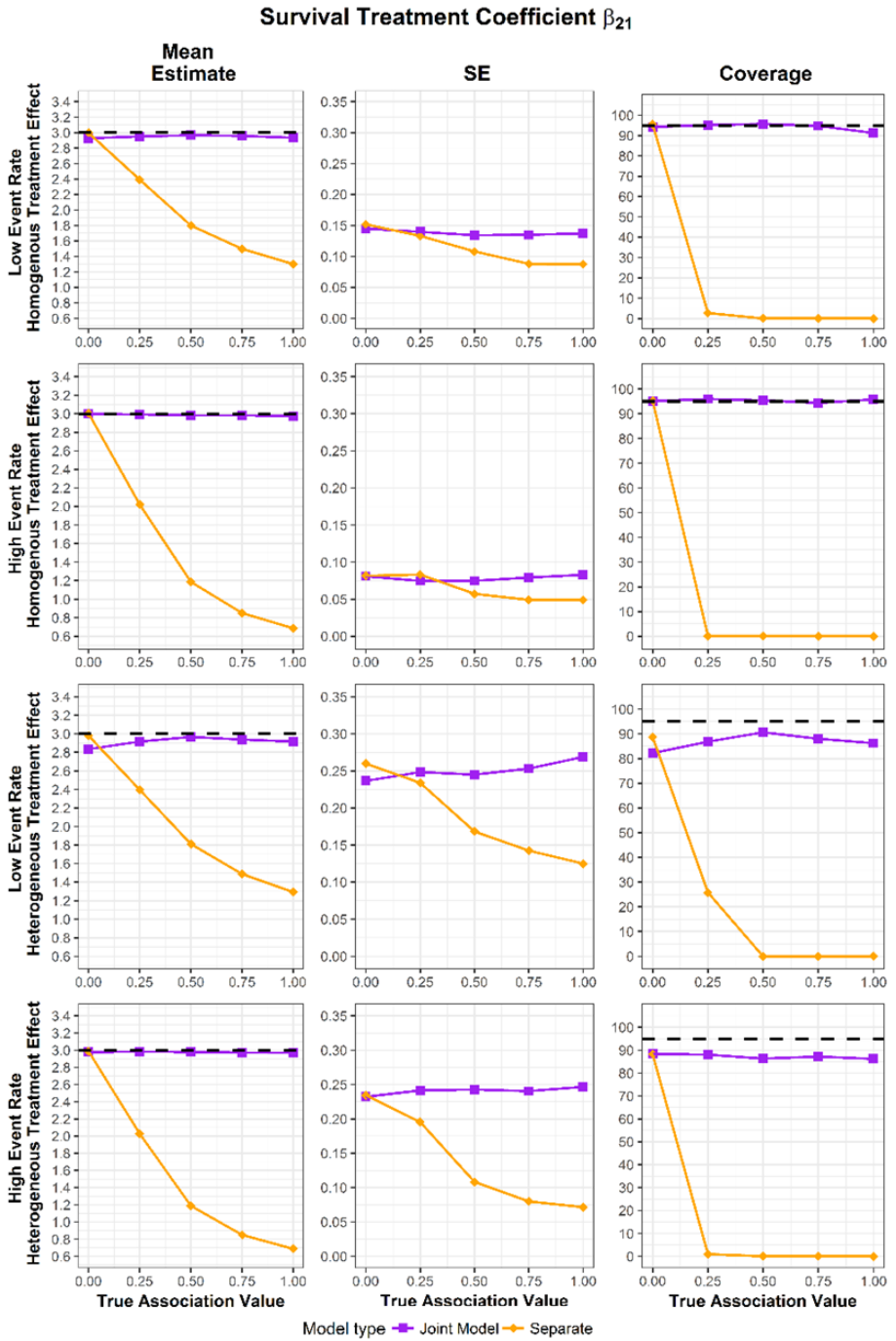
Two stage methods – real data



Two stage methods – simulations (longitudinal treatment effect coefficient)



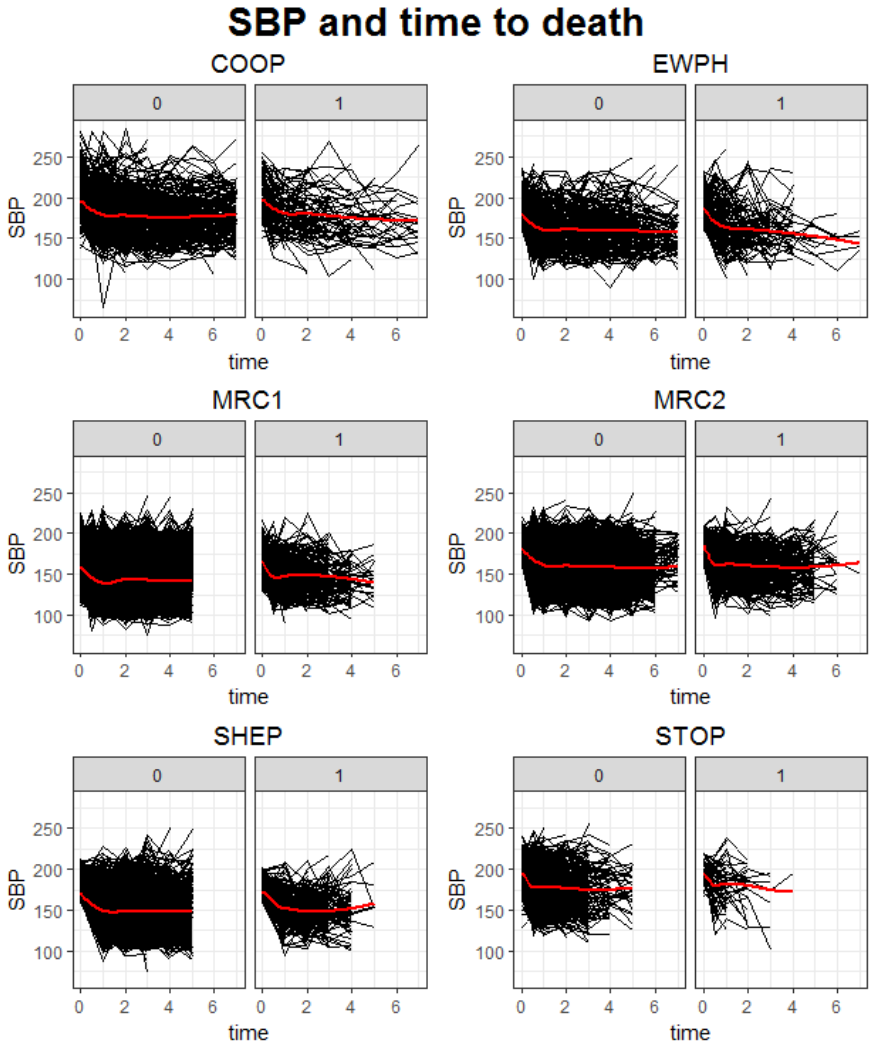
Two stage methods – simulations (time-to-event treatment effect coefficient)



Two stage methods - recommendations

Preliminary work

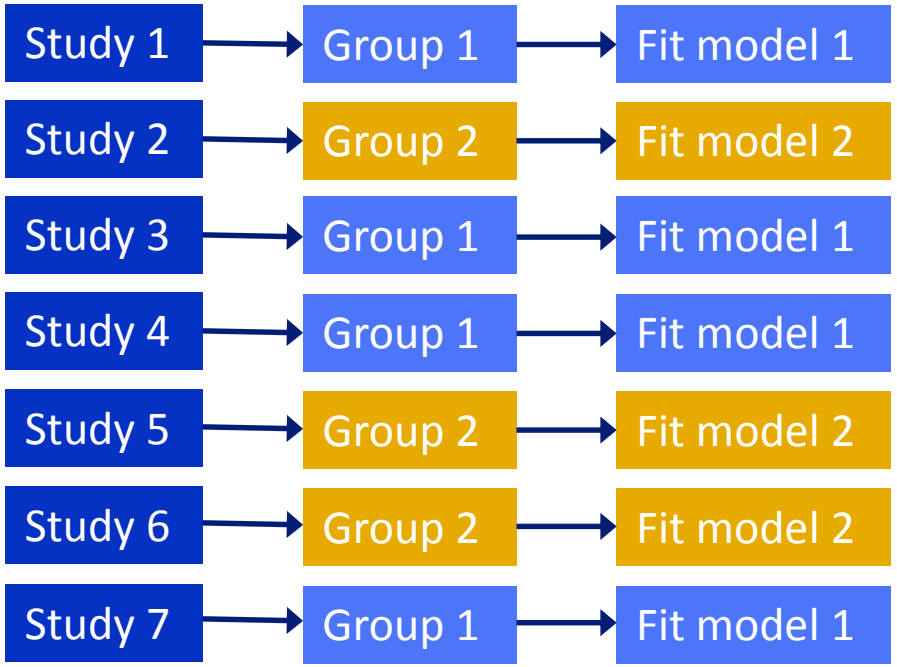
- For each study:
  - Plot longitudinal trajectories separately for those experiencing an event and those censored.
  - Produce Kaplan-Meier plots for e.g. each treatment group
- Use plots to assess whether an association between longitudinal and time-to-event outcomes is feasible
- Use plots and clinical background of the data to select:
  - Longitudinal sub-model
  - Time-to-event sub-model
  - Association structure



Two stage methods - recommendations

First Stage

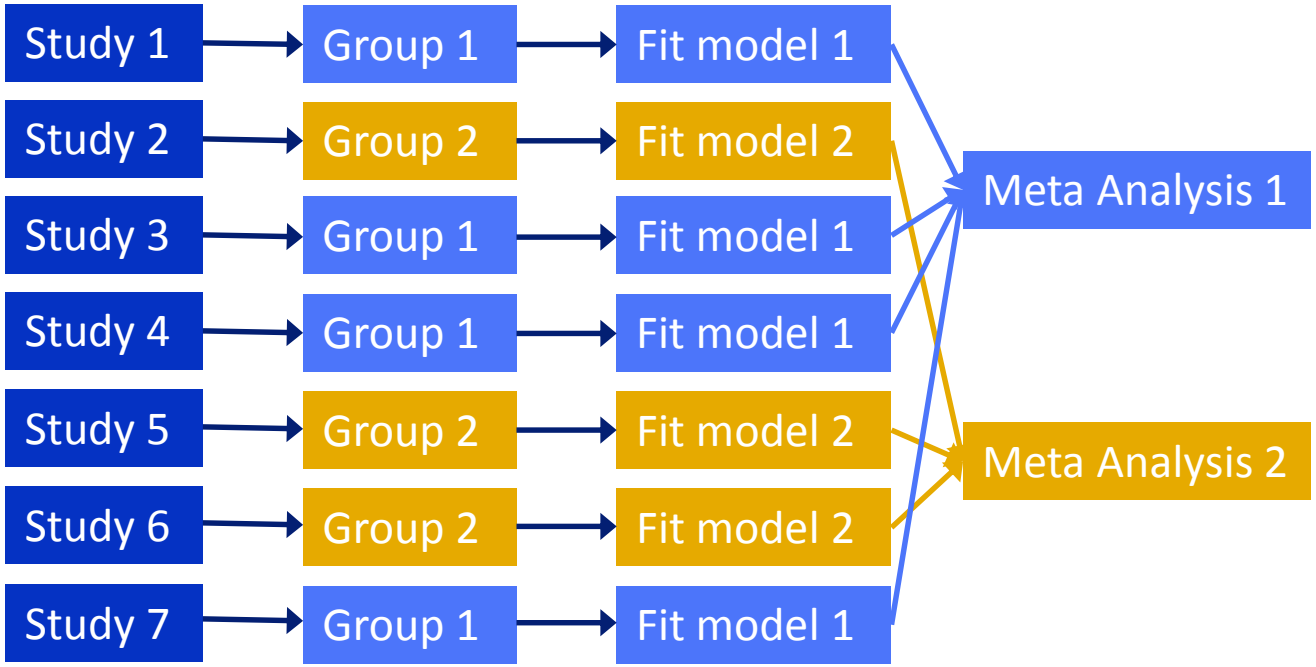
- Group studies such that chosen model structure within each group is identical.
- Within each group, fit identical joint models to data from each study. Model structures can differ between groups.



Two stage methods - recommendations

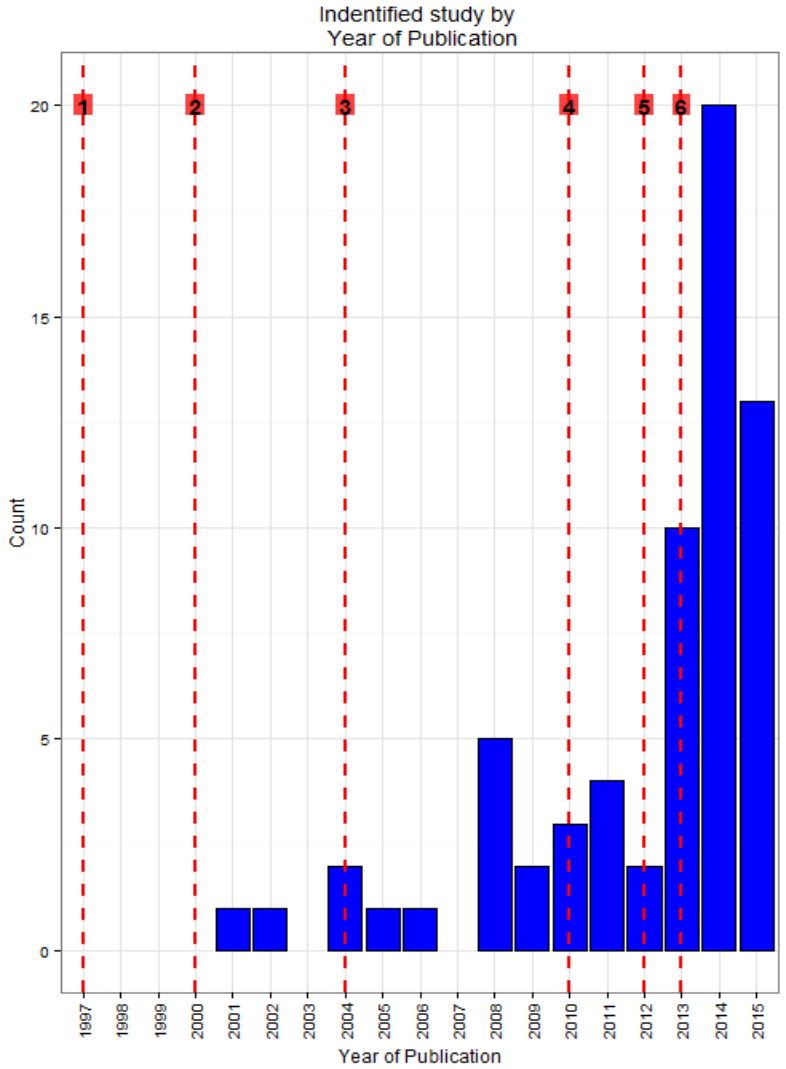
Second Stage

- For each study extract model parameters, precision estimates and sample size
- Pool estimates within groups using standard MA techniques.



Current reporting standard of joint models

Review of standard of published joint data analyses



Review of standard of published joint data analyses

	Longitudinal MA	Time-to-event MA	Association MA
Coefficients reported (%)	45 (69.2)	46 (70.8)	51 (78.5)
Precision reported (%)	44 (67.7)	45 (69.2)	50 (76.9)
Standard Errors reported (%)	22 (33.8)	23 (35.4)	25 (38.5)
Confidence Intervals (CI) reported (%)	30 (46.2)	32 (49.2)	36 (55.4)
Significance level reported (%)	57 (87.7)		
Sample size reported (%)	64 (98.5)		
MA possible given reported information (%)			
All identified studies (N=65)	44 (67.7)	45 (69.2)	50 (76.9)
Studies using joint models to account for dropout (N=22)	18 (81.8)	14 (63.6)	15 (68.2)
Studies using joint models to include time varying covariate in time-to-event sub-model (N=4)	2 (50.0)	3 (75.0)	3 (75.0)

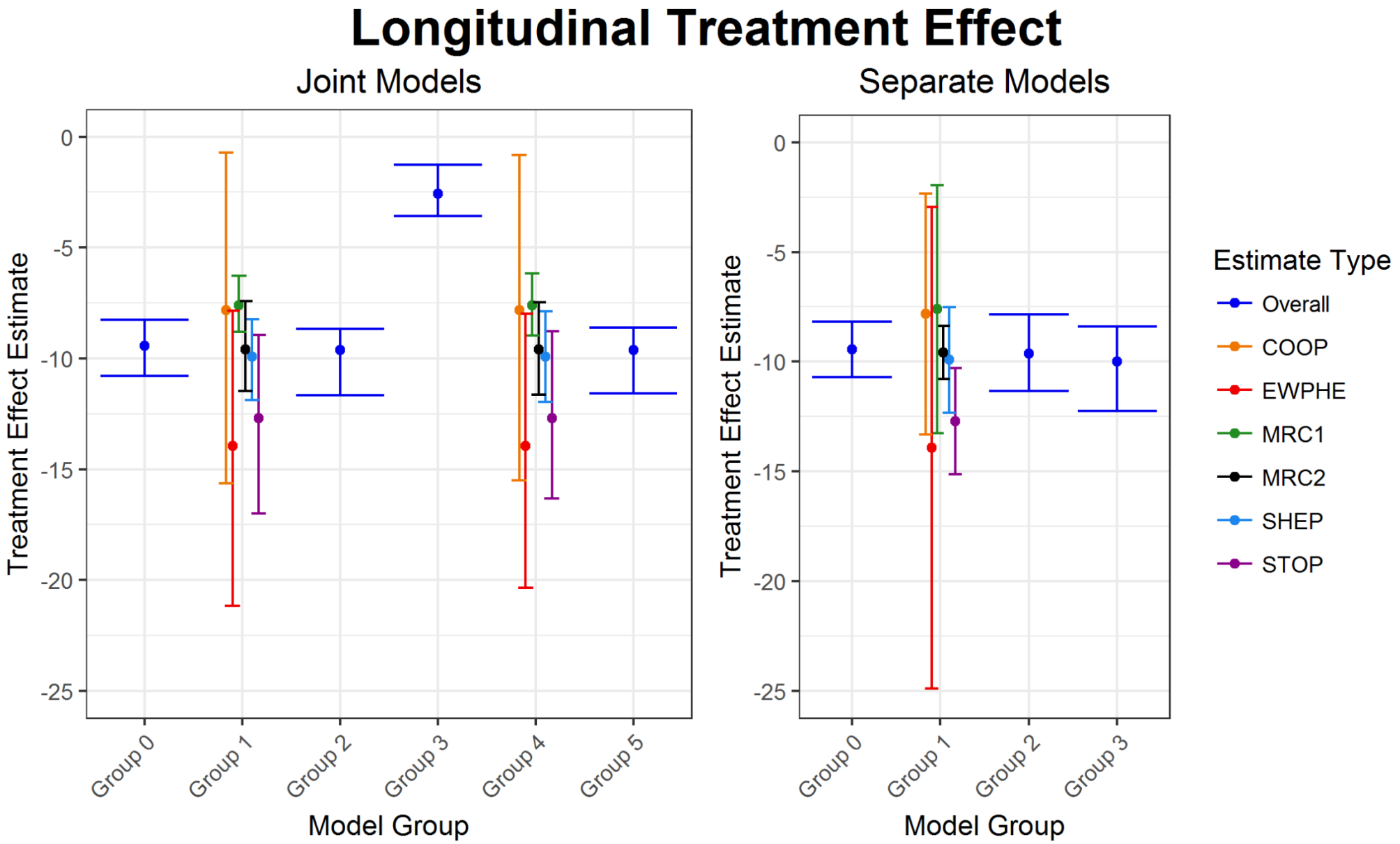
One stage methods

One stage methods - overview

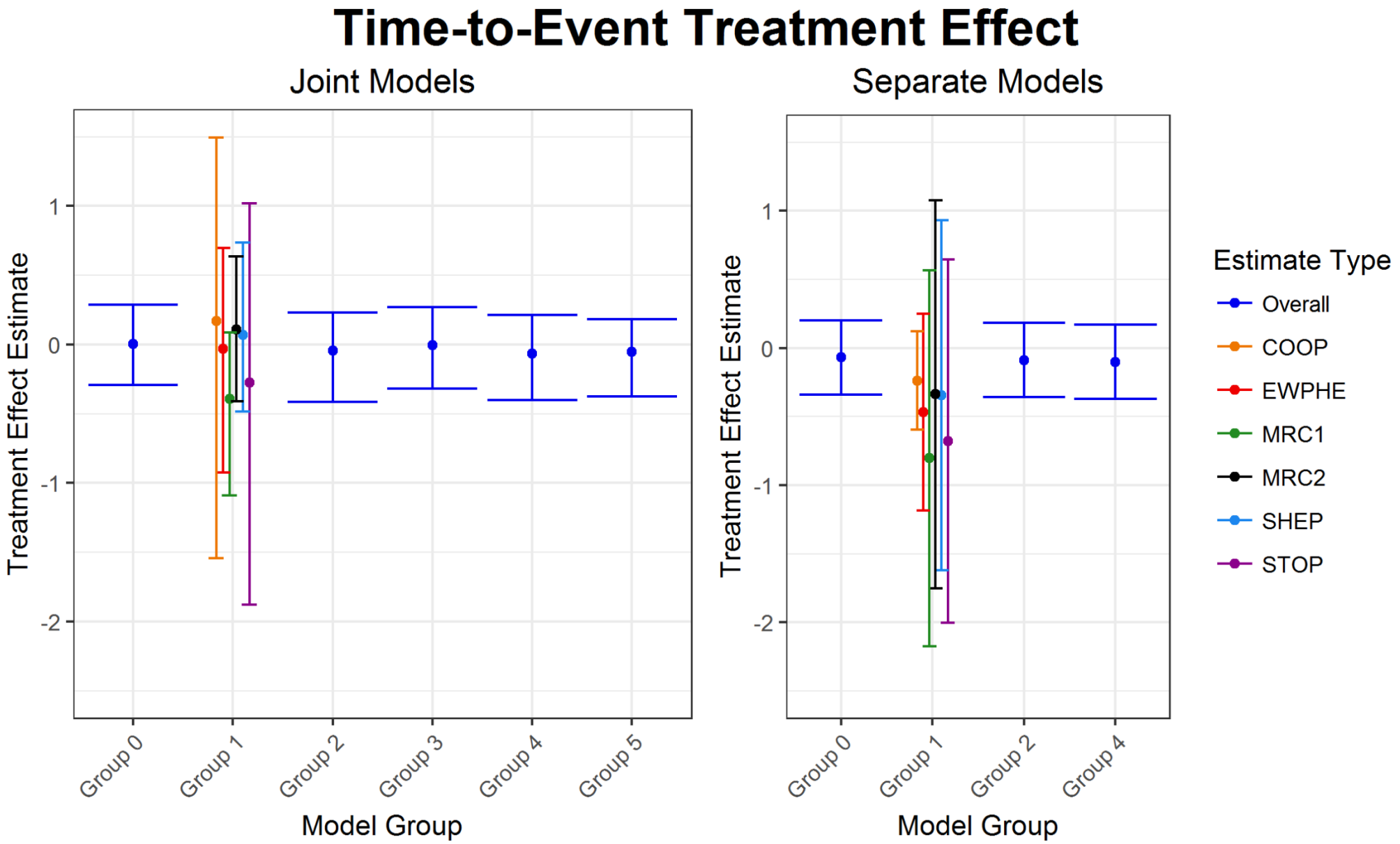
- Longitudinal sub-model: linear mixed effects model with zero mean random effects.
- Time-to-event sub-model: Proportional hazards model with unspecified baseline hazard
- Association structure: random effects proportional association structure
- Aim was not to assess affect of range of covariates, only to assess the different model groups

Group	Method to account for between study heterogeneity
0	Between study heterogeneity ignored
1	Fixed interaction term between treatment and study in each sub-model
2	Fixed study indicator in longitudinal sub-model, study level random treatment effect
3	Study level random intercept and random treatment effect
4	Fixed interaction term between treatment and study in longitudinal sub-model, baseline hazard stratified by study
5	Fixed study indicator in longitudinal sub-model, study level random treatment effect, baseline hazard stratified by study

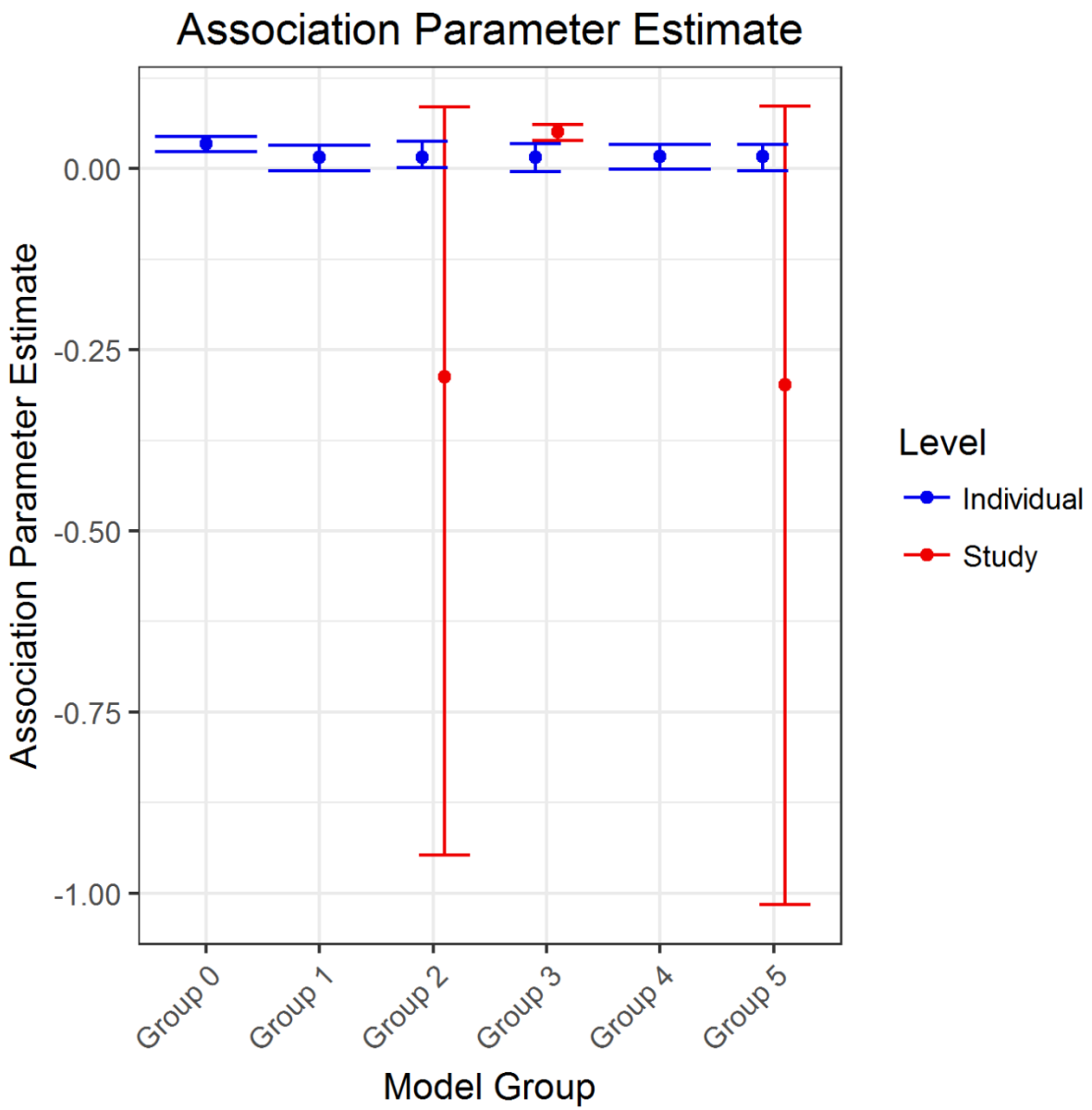
One stage methods – real data



One stage methods – real data



One stage methods – real data



One stage methods – simulation work

- Ongoing
- Scenarios designed to investigate behaviour of different model groups under e.g.
  - Varying association levels
  - Low (~25%) or high (~75%) event rate
  - Varying numbers of included studies

Software – joineRmeta package

- Currently available from GitHub - https://github.com/mesudell/joineRmeta/

- Functions to:
  - Easily plot multi-study joint data
  - To automatically extract and meta-analyse specified model parameters from supplied joint model fits
  - Model three level joint data allowing for
    - Random effects only association structure
    - Individual level and study level random effects
    - Un-stratified or stratified baseline hazard
  - Simulation of multi-study joint data

Discussion

- One stage methods
  - Use of study level random effects may be unwise unless number of studies is over a certain threshold
  - Interaction terms between covariates and study membership would quickly become cumbersome with large numbers of studies
  - Allows in depth investigation of between study heterogeneity
- Two stage methods
  - Faster than one stage methods
  - Increasing difference between coefficient estimates in separate time-to-event models compared to joint models as association increases in magnitude
  - Multivariate meta-analysis of results rather than separate meta-analyses of each coefficient might be beneficial – future research
- Time commitment to bootstrap models is a concern

Conclusions

- Care must be taken during two stage meta-analyses of joint data to pool only parameters with comparable interpretations
- Current reporting of joint models may hamper an aggregate data meta-analysis of joint data given current reporting standards
- A variety of methods exist to model multi-study joint data in a one stage analyses, however some may not be appropriate unless the number of studies in the meta-analysis is over a given threshold

Thank you for listening.
Any questions?